

# ANNALES DE L'I. H. P., SECTION B

MONIQUE JACCOTTET

RICHARD TOMASSONE

## **Méthodes d'analyse factorielle en théorie des plans d'expérience**

*Annales de l'I. H. P., section B*, tome 12, n° 3 (1976), p. 233-256

[http://www.numdam.org/item?id=AIHPB\\_1976\\_\\_12\\_3\\_233\\_0](http://www.numdam.org/item?id=AIHPB_1976__12_3_233_0)

© Gauthier-Villars, 1976, tous droits réservés.

L'accès aux archives de la revue « Annales de l'I. H. P., section B » (<http://www.elsevier.com/locate/anihpb>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

## Méthodes d'analyse factorielle en théorie des plans d'expérience

par

**Monique JACCOTTET (\*) et Richard TOMASSONE (\*\*)**

(\*) Équipe de statistique. Université Paris-Sud 91405 Orsay  
adresse actuelle : Département de Mathématiques,  
École Polytechnique Fédérale de Lausanne, CH1007 Lausanne

(\*\*) I. N. R. A., Laboratoire de Biométrie du C. N. R. Z., 78350 Jouy-en-Josas

---

**RÉSUMÉ.** — L'analyse des correspondances et l'analyse des corrélations canoniques sont utilisées pour étudier la structure d'un plan d'expérience à deux facteurs. Les propriétés d'équilibre du plan, ainsi que son efficacité sont mises en relation avec des résultats connus de l'analyse factorielle.

**SUMMARY.** — Correspondence analysis and canonical correlations analysis are used to study the structure of a two-way design of experiments. The properties of equilibrium and the efficiency of the design are linked with known results of factor analysis.

---

### INTRODUCTION

Il ne s'agit pas ici — comme on le fait du reste depuis longtemps — d'appliquer une analyse de variance multivariée ou une analyse factorielle à des données résultant d'un plan d'expérience, mais d'utiliser des méthodes d'analyse factorielle pour l'étude de la structure du plan, indépendamment du matériel expérimental.

Le tableau des données est alors la matrice d'incidence du plan, c'est-à-dire la table de contingence entre blocs et traitements. Cette matrice

définit entièrement le dispositif expérimental : elle détermine la précision et l'efficacité pour estimer (sous l'hypothèse du modèle linéaire additif) les paramètres et les contrastes ; elle fixe également, à un facteur de proportionnalité près, la matrice de covariance des traitements et des blocs. Le plan d'expérience est donc considéré comme *l'expression des relations de dépendance entre blocs et traitements*.

L'analyse des correspondances est une méthode puissante pour l'étude des tables de contingence. Nous verrons que les caractéristiques d'un dispositif expérimental s'interprètent aisément en termes d'inertie et d'axes d'élongation du nuage des individus ou des variables, lorsqu'on définit les individus comme les traitements et les variables comme les blocs.

## 1. ANALYSE DES CORRESPONDANCES SUR UNE MATRICE D'INCIDENCE

### 1.1. Notations relatives aux plans d'expérience

Nous donnons ci-dessous quelques notations classiques dans le domaine d'étude des plans d'expérience, puis quelques résultats non moins classiques dans celui de l'étude de l'analyse des correspondances ; le seul intérêt de cette juxtaposition réside dans le fait que les deux domaines sont étudiés, dans la plupart des cas, dans des « milieux scientifiques » différents.

Soit la matrice d'incidence du plan  $n = (n_{ij})$ ,  $i = 1 \dots t$ ;  $j = 1 \dots b$ . Le nombre de répétitions de chaque traitement  $i$ ,  $r_i = \sum_j n_{ij}$  est regroupé dans

un vecteur  $r' = (r_1, r_2, \dots, r_t)$ ; nous utiliserons aussi la matrice diagonale  $r^\delta$ , dont les éléments diagonaux sont les composantes de  $r$ . Notons aussi que  $r^{\alpha\delta} = (r^\delta)^\alpha$  pour  $\alpha \in \mathbb{R}$ , identité que nous appliquerons plus loin pour  $\alpha = -1$  et  $\alpha = -\frac{1}{2}$ . Nous avons besoin de notations analogues pour le nombre

d'observations dans chaque bloc  $j$ ,  $k_j = \sum_i n_{ij}$ ,  $k' = (k_1, k_2, \dots, k_b)$  et  $k^\delta$ .

Le nombre total d'unités expérimentales est  $N = \sum_i r_i = \sum_j k_j$ .

Enfin, on peut toujours décomposer la matrice d'incidence de la manière suivante :

$$n = rk'/N + A$$

où  $A$ , matrice de même dimension que  $n$ , est appelée *matrice de non-orthogonalité*. Pour un plan orthogonal, c'est-à-dire pour un plan où les estimateurs des traitements sont indépendants de ceux des blocs,  $A$  est une matrice formée de zéros; en terme probabiliste la loi composée  $n_{ij}/N$  est le produit des lois marginales  $r_i/N$  et  $k_j/N$ .

**1.2. Résultats de l'analyse des correspondances.**

Si nous passons maintenant à la terminologie et aux résultats de l'analyse des correspondances nous allons découvrir une similitude frappante; afin de la mieux saisir et de rappeler quelques résultats qui nous seront utiles par la suite nous allons les donner, sans démonstration aucune, avec les notations du paragraphe précédent.

1.2.1. De manière générale, si  $R^b$  est l'espace canonique à  $b$  dimensions ( $b$  étant le nombre de blocs), on définit pour chaque bloc la « probabilité conditionnelle d'apparition du bloc  $j$  étant donné le traitement  $i$  », c'est-à-dire le nombre

$$(n_{ij}/N)/(r_i/N) = n_{ij}/r_i \quad i = 1, \dots, t \tag{1}$$

qui représente en analyse des correspondances la projection du traitement  $i$  sur le  $j^e$  axe de  $R^b$ . Considérant l'ensemble des blocs, on appelle *profil du traitement  $i$*  le point de  $R^b$  défini par :

$$(n_{ij}/r_i, \quad j = 1, \dots, b).$$

L'espace  $R^b$ , appelé *espace des blocs*, est muni de la métrique

$$Nk^{-\delta}$$

ce qui revient à pondérer par les « lois marginales » des blocs.

La distance entre deux traitements est alors définie par :

$$d^2(i, i') = N \sum_j k_j^{-1} \left( \frac{n_{ij}}{r_i} - \frac{n_{i'j}}{r_{i'}} \right)^2 \tag{2}$$

$$\frac{n_{ij}}{r_i} - \frac{n_{i'j}}{r_{i'}} = \frac{r_i k_j}{Nr_i} + \frac{a_{ij}}{r_i} - \left( \frac{r_i \cdot k_j}{Nr_{i'}} + \frac{a_{i'j}}{r_{i'}} \right) = \frac{a_{ij}}{r_i} - \frac{a_{i'j}}{r_{i'}}$$

La distance entre deux traitements n'est donc pas changée, s'ils sont représentés par les points  $\left( \frac{a_{ij}}{r_i}, j = 1, \dots, b \right)$ .

Effectuons de plus le changement d'échelle suivant sur chaque axe :

$$Nk_j^{-1} \quad j = 1, \dots, b$$

Le traitement  $i$  est maintenant représenté par le point :

$$\left( \left( \sqrt{N} \frac{a_{ij}}{r_i \sqrt{k_j}} \right), \quad j = 1, \dots, b \right) \quad (3)$$

et la distance entre deux traitements :

$$d^2(i, i') = \sum_j \left( \frac{a_{ij} \sqrt{N}}{r_i \sqrt{k_j}} - \frac{a_{i'j} \sqrt{N}}{r_{i'} \sqrt{k_j}} \right)^2 = N \sum_j k_j^{-1} \left( \frac{a_{ij}}{r_i} - \frac{a_{i'j}}{r_{i'}} \right)^2$$

reste inchangée.

L'ensemble des points

$$\left( \frac{a_{ij} \sqrt{N}}{r_i \sqrt{k_j}}, \quad j = 1, \dots, b \right), \quad i = 1, \dots, t$$

est appelé le nuage des traitements dans l'espace des blocs.

*Remarques.* — 1) Comme les  $a_{ij}$  sont les termes de la matrice de non-orthogonalité, le nuage des traitements est réduit à l'origine dans le cas du plan orthogonal.

2) Le nuage des traitements appartient à l'hyperplan d'équation

$$\sum_j k_j^{\frac{1}{2}} x_j = 0 \quad (4)$$

3) L'inertie du nuage par rapport à l'origine pour l'ensemble de poids  $(r_i/N, i = 1, \dots, t)$ , fréquences relatives des traitements, est définie par :

$$\sum_i \frac{r_i}{N} \sum_j \left( \frac{a_{ij} \sqrt{N}}{r_i \sqrt{k_j}} \right)^2 = \sum_{i,j} \frac{a_{ij}^2}{r_i k_j} = \sum_{i,j} \frac{(n_{ij} - r_i k_j / N)^2}{r_i k_j N} \quad (5)$$

est analogue à un  $\chi^2$ , elle fournit une mesure de dépendance entre blocs et traitements (Lebart et Fenelon, 1975).

Plus généralement, la matrice d'inertie du nuage est, pour le système de poids  $(r_i/N)$  :

$$V_{jj'} = \sum_{i=1}^t \frac{r_i}{N} \left( \frac{N^{\frac{1}{2}} a_{ij}}{r_i k_j^{\frac{1}{2}}} \cdot \frac{N^{\frac{1}{2}} a_{ij'}}{r_i k_{j'}^{\frac{1}{2}}} \right) = \sum_{i=1}^t \frac{a_{ij} a_{ij'}}{r_i k_j^{\frac{1}{2}} k_{j'}^{\frac{1}{2}}} = \frac{1}{(k_j k_{j'})^{\frac{1}{2}}} \sum_{i=1}^t \frac{a_{ij} a_{ij'}}{r_i}$$

soit

$$\mathbf{V} = \mathbf{k}^{-\delta/2} \mathbf{A}' \mathbf{r}^{-\delta} \mathbf{A} \mathbf{k}^{-\delta/2} \tag{6}$$

Posons  $\xi = \mathbf{r}^{-\delta/2} \mathbf{A} \mathbf{k}^{-\delta/2}$ , alors

$$\mathbf{V} = \xi' \xi \quad \text{matrice d'inertie dans l'espace des blocs.}$$

1.2.2. On peut considérer, de manière symétrique, un nuage des blocs dans l'espace des traitements :

$$\left( \frac{a_{ij} N^{\frac{1}{2}}}{k_j r_i^{\frac{1}{2}}}, \quad i = 1, \dots, t, \quad j = 1, \dots, b \right) \tag{7}$$

dont la matrice d'inertie, pour le système de poids  $(k_j/N)$  s'écrit :

$$\mathbf{H} = \xi \xi' \quad \text{matrice d'inertie dans l'espace des traitements.}$$

Cette matrice est liée d'une manière simple à la matrice  $\mathbf{\Omega}$ , qui est définie comme la matrice des variances et covariances (à un facteur près, égal à la variance des résidus) des estimateurs des paramètres des traitements (Pearce, 1970)

$$\mathbf{\Omega}^{-1} = \mathbf{r}^{\delta/2} (\mathbf{I} - \mathbf{H}) \mathbf{r}^{\delta/2}$$

### 1.3. Efficacité d'un plan d'expérience.

L'efficacité d'un plan d'expérience, qui est en quelque sorte une mesure *a priori* de son « rendement » (Kempthorne, 1952 ; Cochran et Cox, 1968), s'interprète alors très simplement à l'aide des valeurs propres  $\lambda_h$  de  $\mathbf{H}$  (qui sont aussi celles de  $\mathbf{V}$ ) ; elles donnent une expression de l'efficacité moyenne du plan pour estimer les traitements :

$$\varepsilon = \frac{t}{t + E} \quad \text{où} \quad E = \sum_h \frac{\lambda_h}{1 - \lambda_h}.$$

On peut trouver de même l'efficacité moyenne du plan dual, pour estimer les blocs :

$$\tilde{\varepsilon} = \frac{b}{b + E}$$

#### 1.4. La détermination des axes d'inertie en analyse des correspondances.

Une fois trouvées les matrices  $\mathbf{V}$  et  $\mathbf{H}$ , la technique de l'analyse des correspondances consiste en la recherche des axes d'inertie du nuage des traitements (respectivement des blocs), qui sont les directions des vecteurs propres de  $\mathbf{V}$  (respectivement de  $\mathbf{H}$ ).

Si  $\mathbf{u}_h$  est vecteur propre de  $\mathbf{H}$  :

$$\lambda_h \mathbf{u}_h = \mathbf{H} \mathbf{u}_h = \boldsymbol{\xi} \boldsymbol{\xi}' \mathbf{u}_h$$

il donne la direction du  $h^e$  axe d'inertie du nuage des blocs, et l'axe d'inertie correspondant à la même valeur propre  $\lambda_h$ , supposée simple, dans l'espace des blocs est donné par le  $h^e$  vecteur propre normé  $\tilde{\mathbf{u}}_h$  de  $\mathbf{V}$ , et peut s'exprimer en fonction de  $\mathbf{u}_h$  :

$$\begin{aligned} \lambda_h \tilde{\mathbf{u}}_h &= \mathbf{V} \tilde{\mathbf{u}}_h = \boldsymbol{\xi}' \boldsymbol{\xi} \tilde{\mathbf{u}}_h \\ \lambda_h \boldsymbol{\xi}' \mathbf{u}_h &= \boldsymbol{\xi}' (\boldsymbol{\xi} \boldsymbol{\xi}') \mathbf{u}_h = \mathbf{V} \boldsymbol{\xi}' \mathbf{u}_h. \end{aligned}$$

Si  $\mathbf{u}_h' \mathbf{u}_h = 1$  :

$$\mathbf{u}_h' \boldsymbol{\xi} \boldsymbol{\xi}' \mathbf{u}_h = \lambda_h \mathbf{u}_h' \mathbf{u}_h = \lambda_h \Rightarrow \tilde{\mathbf{u}}_h = (\lambda_h)^{-\frac{1}{2}} \boldsymbol{\xi}' \mathbf{u}_h \quad (8)$$

Cette relation permet une représentation simultanée du nuage des traitements et de celui des blocs.

Cherchons, dans l'espace des blocs, la composante  $f_h(i)$  du traitement  $i$  sur le  $h^{\text{ème}}$  axe d'inertie :

$$f_h(i) = \sum_{j=1}^b \frac{N^{\frac{1}{2}} a_{ij}}{r_i k_j^{\frac{1}{2}}} \tilde{u}_{hj} \quad (9)$$

(où  $\tilde{u}_{hj}$  est la  $j^{\text{ème}}$  composante du vecteur  $\tilde{\mathbf{u}}_h$  dans la base canonique)

$$(f_h(i), \quad i = 1, \dots, t) = N^{\frac{1}{2}} \mathbf{r}^{-\delta} \mathbf{A} \mathbf{k}^{-\delta/2} \tilde{\mathbf{u}}_h \quad (9')$$

représente l'ensemble des  $h^{\text{ème}}$  composantes du nuage. Les  $t$  quantités  $f_h(i) \mathbf{u}_h$  sont alors les projections des  $t$  traitements sur le  $h^{\text{ème}}$  axe ; par un abus de langage, qui n'introduit aucune ambiguïté, nous appellerons désormais  $f_h(i)$  la projection du traitement  $i$  sur l'axe  $h$ .

Son image dans l'espace des traitements vaut :

$$\begin{aligned} (f_h(i), \quad i = 1, \dots, t) &= N^{\frac{1}{2}} \mathbf{r}^{-\delta} \mathbf{A} \mathbf{k}^{-\delta/2} (\lambda_h^{-\frac{1}{2}} \mathbf{k}^{-\delta/2} \mathbf{A}' \mathbf{r}^{-\delta/2}) \mathbf{u}_h \\ &= \left( \frac{N}{\lambda_h} \right)^{\frac{1}{2}} \mathbf{r}^{-\delta/2} \mathbf{H} \mathbf{u}_h = (N \lambda_h)^{\frac{1}{2}} \mathbf{r}^{-\delta/2} \mathbf{u}_h \quad (10) \end{aligned}$$

On définit alors le vecteur

$$\mathbf{t}_h = \mathbf{N}^{\frac{1}{2}} \mathbf{r}^{-\delta/2} \mathbf{u}_h \tag{11}$$

appelé *h<sup>e</sup> axe factoriel* dans l'espace des traitements ; cet axe est le sous-espace engendré par le vecteur.

La projection du traitement *i* sur cet axe vaut :

$$f_h(i) = \lambda_h^{\frac{1}{2}} t_{hi} = \lambda_h^{\frac{1}{2}} \delta'_i t_h$$

où  $\delta'_i = (0 \dots 0 \underset{i^e}{1} 0 \dots 0)$  ( $1 \times t$ ).

$f_h(i)$  est donc, au facteur  $\lambda_h^{\frac{1}{2}}$  près, la projection du *i<sup>e</sup>* vecteur de la base canonique sur  $\mathbf{t}_h$ .

A chaque traitement correspond un vecteur de la base canonique.

La projection du bloc *j* sur  $\mathbf{u}_h$  vaut :

$$\sum_{i=1}^t \frac{a_{ij} \sqrt{\mathbf{N}}}{k_j \sqrt{r_i}} \mathbf{u}_{hi} = \sum_{i=1}^t \frac{a_{ij}}{k_j} t_{hi} \tag{12}$$

En résumé, posons  $f_h(\mathbf{I})$  pour l'ensemble des *h<sup>e</sup>* projections des traitements et  $f_h(\mathbf{J})$  pour les blocs, alors :

$$\begin{aligned} f_h(\mathbf{I}) &= \lambda_h^{\frac{1}{2}} \mathbf{t}_h \\ f_h(\mathbf{J}) &= \mathbf{k}^{-\delta} \mathbf{A}' \mathbf{t}_h \end{aligned} \tag{13}$$

(On pourrait trouver des relations analogues dans l'espace des blocs ; donnons seulement une équation liant  $\mathbf{t}_h$  et  $\tilde{\mathbf{t}}_h$  :)

$$\begin{aligned} \tilde{\mathbf{t}}_h &= \mathbf{N}^{\frac{1}{2}} \mathbf{k}^{-\delta/2} \tilde{\mathbf{u}}_h = \mathbf{N}^{\frac{1}{2}} \mathbf{k}^{-\delta/2} (\lambda_h)^{-\frac{1}{2}} \boldsymbol{\xi}' \mathbf{u}_h \\ &= (\mathbf{N}/\lambda_h)^{\frac{1}{2}} \mathbf{k}^{-\delta/2} \mathbf{k}^{-\delta/2} \mathbf{A}' \mathbf{r}^{-\delta/2} \mathbf{u}_h = (\lambda_h)^{-\frac{1}{2}} \mathbf{k}^{-\delta} \mathbf{A}' \mathbf{t}_h \end{aligned}$$

$$\tilde{\mathbf{t}}_h = \lambda_h^{-1/2} \mathbf{k}^{-\delta} \mathbf{A}' \mathbf{t}_h \tag{14}$$

## 2. INTERPRÉTATION DES AXES FACTORIELS DANS LA THÉORIE DES PLANS D'EXPÉRIENCE

Calinski (1971, 1972) a montré l'utilité de la matrice

$$\mathbf{M}_0 = \mathbf{r}^{-\delta} \mathbf{A} \mathbf{k}^{-\delta} \mathbf{A}'$$



pour déterminer les propriétés de J-équilibre du plan (\*); il a établi:

$$\begin{aligned}\Omega^{-1} &= r^\delta(\mathbf{I} - \mathbf{M}_0) \\ \Omega &= \left( \mathbf{I} + \sum_{l=1}^{\infty} \mathbf{M}_0^l \right) r^{-\delta}\end{aligned}$$

DÉFINITION. — On appelle  $r$ -contraste un vecteur de  $\mathbf{R}^t$  tel que son produit scalaire avec  $r$  soit nul.

Résultats. — Les vecteurs propres de  $\mathbf{M}_0$  associés aux valeurs propres non nulles sont des  $r$ -contrastes.

Soit  $t$  un vecteur propre de  $\mathbf{M}_0$ , alors  $c = r^{-\delta}t$  est un contraste, pour lequel la précision du plan et l'efficacité sont les mêmes (Pearce *et al.*, 1974).

Les axes factoriels  $t_h$  sont précisément les vecteurs propres associés aux valeurs propres non nulles de  $\mathbf{M}_0$ . Le plan est donc, par définition, J-équilibré pour les axes factoriels.

En effet :

$$\begin{aligned}\mathbf{H}u_h &= \lambda_h u_h \\ \Leftrightarrow r^{-\delta/2} \mathbf{A} \mathbf{k}^{-\delta} \mathbf{A}' r^{-\delta/2} u_h &= \mathbf{N}^{-\frac{1}{2}} r^{-\delta/2} \mathbf{A} \mathbf{k}^{-\delta} \mathbf{A}' t_h = \mathbf{N}^{-\frac{1}{2}} \lambda_h r^{\delta/2} t_h \quad (\text{éq. (11)}) \\ \Leftrightarrow r^{-\delta} \mathbf{A} \mathbf{k}^{-\delta} \mathbf{A}' t_h &= \mathbf{M}_0 t_h = \lambda_h t_h \quad (15)\end{aligned}$$

Expression de  $\Omega$  en fonction des axes factoriels. — On suppose le plan connexe, i. e.  $\Omega$  de rang  $t$ .

$\mathbf{I} - \mathbf{M}_0$  est alors de rang  $t$ , puisque

$$\Omega^{-1} = r^\delta(\mathbf{I} - \mathbf{M}_0)$$

Remarque. —  $\mathbf{I} - \mathbf{M}_0$  inversible et définie positive

$$\Rightarrow \lambda_h < 1 \quad \forall h.$$

Soit  $t_0$  le rang de  $\mathbf{M}_0$ , et soit  $\mathbf{T}$  la matrice de vecteurs propres de  $\mathbf{I} - \mathbf{M}_0$  disposés en colonnes

$$t_0 \leq t - 1$$

puisque les axes factoriels sont dans un hyperplan de  $\mathbf{R}^t$ .

La valeur propre 0 de  $\mathbf{M}_0$  est de multiplicité  $t - t_0$ . On complète la base de vecteurs propres  $\{\mathbf{N}^{-\frac{1}{2}} t_h, h = 1, \dots, t_0\}$  par  $t - t_0 - 1$   $r$ -contrastes linéairement indépendants et par le vecteur  $\mathbf{N}^{-\frac{1}{2}} \mathbf{1}_t$ , de telle sorte que les relations

$$\begin{aligned}\mathbf{N}^{-1} t'_h r^\delta t_h &= 1 \\ t'_{h_1} r^\delta t_{h_2} &= 0\end{aligned}$$

(\*) Rappelons qu'un plan est équilibré au sens de l'efficacité ou J-équilibré si les contrastes sont estimés avec la même efficacité (JONES, 1959).



Plan orthogonal :

$$\mathbf{T}'\mathbf{r}^\delta(\mathbf{r}^{-\delta})\mathbf{r}^\delta\mathbf{T} = \mathbf{T}'\mathbf{r}^\delta\mathbf{T} = \mathbf{I}_t \Rightarrow \mathbf{V}'\mathbf{r}^{-\delta}\mathbf{V} = \mathbf{I}_{t-1}.$$

CONCLUSIONS. — 1) Les  $t - t_0 - 1$  contrastes appartenant au noyau de  $\mathbf{M}_0$  sont estimés avec l'efficacité (et la précision) maximum.

On sait, par l'analyse des correspondances, que plus les valeurs propres  $\lambda_h$  sont proches de 0, moins les deux ensembles (blocs et traitements) sont corrélés, et plus l'efficacité du plan augmente.

2)  $\Omega$  est une fonction des  $r$ -contrastes de base et du facteur d'efficacité ; elle ne dépend donc pas de la répartition des unités expérimentales dans les blocs.

3) Si  $t_0 = rg\mathbf{M}_0 = t - 1$ , la représentation des traitements donnée par l'analyse des correspondances

$$f_h(\mathbf{I}) = (\lambda_h)^{\frac{1}{2}} \cdot \mathbf{t}_h$$

est donc équivalente à la donnée de la matrice  $\Omega$ .

4) Si  $t_1 = rg\mathbf{M}_0 < t - 1$ ,  $\Omega$  peut s'exprimer en fonction de  $r$  et des facteurs  $\mathbf{t}_h$ ,  $h \leq t_1$  :

$$\begin{aligned} \mathbf{T}\mathbf{T}' = \mathbf{r}^{-\delta} &= \mathbf{N}^{-1} \sum_{h=1}^t \mathbf{t}_h \mathbf{t}'_h = \mathbf{N}^{-1} \sum_1^{t_1} \mathbf{t}_h \mathbf{t}'_h + \mathbf{N}^{-1} \sum_{t_1+1}^t \mathbf{t}_h \mathbf{t}'_h \\ &\Rightarrow \mathbf{N}^{-1} \sum_{t_1+1}^t \mathbf{t}_h \mathbf{t}'_h = \mathbf{r}^{-\delta} - \mathbf{N}^{-1} \sum_1^{t_1} \mathbf{t}_h \mathbf{t}'_h \\ \Omega &= \mathbf{T}(\mathbf{1} - \lambda)^{-\delta} \mathbf{T}' = \mathbf{N}^{-1} \sum_1^{t_1} (1 - \lambda_h)^{-1} \mathbf{t}_h \mathbf{t}'_h + \mathbf{N}^{-1} \sum_{t_1+1}^t \mathbf{t}_h \mathbf{t}'_h \end{aligned}$$

$$\boxed{\Omega = \mathbf{r}^{-\delta} + \mathbf{N}^{-1} \sum_1^{t_1} \frac{\lambda_h}{1 - \lambda_h} \mathbf{t}_h \mathbf{t}'_h} \quad (18')$$

### 3. CAS DES DISTANCES NULLES

#### 3.1. Traitements proportionnels

Les traitements  $i$  et  $i'$  sont dits proportionnels si

$$n_{ij} = \theta n_{i'j} \quad (\theta = r_i r_{i'}^{-1})$$

ou

$$\frac{a_{ij}}{r_i} = \frac{a_{i'j}}{r_{i'}} \quad (20)$$

Les projections des traitements  $i$  et  $i'$  en analyse des correspondances sur les axes factoriels  $\tilde{t}_h$  sont :

$$f_h(i) = \lambda_h^{\frac{1}{2}} t_{hi} = N^{\frac{1}{2}} \sum_j \frac{a_{ij}}{r_i} \tilde{t}_{hj}$$

$$f_h(i') = \lambda_h^{\frac{1}{2}} t_{hi'} = N^{\frac{1}{2}} \sum_j \frac{a_{i'j}}{r_{i'}} \tilde{t}_{hj}$$

(20) est équivalente à  $f_h(i) = f_h(i')$  (21)

Les projections sur les axes factoriels de deux traitements proportionnels sont confondues.

*L'identification de deux traitements proportionnels ne change pas la représentation graphique donnée par l'analyse des correspondances.*

Cependant, la condition pour que la précision soit invariante lors de l'identification de deux traitements est plus souple.

En effet :  $t_{hi} = t_{hi'}$  pour tout  $h$  signifie que la matrice :

$$r^{-\delta/2} \mathbf{H} r^{-\delta/2} = N^{-1} \sum_{h=1}^{t_0} \lambda_h t_h t_h' = \Delta$$

a deux lignes égales

$$\Omega^{-1} = r^\delta - r^{\delta/2} \mathbf{H} r^{\delta/2}.$$

Posons

$$\Omega^{ij} = (\Omega^{-1})_{ij}.$$

On en déduit :

$$\left. \begin{aligned} r_2 \Omega^{1j} &= r_1 \Omega^{2j} & j = 3, \dots, t \\ r_2 (\Omega^{11} + \Omega^{12}) &= r_1 (\Omega^{22} + \Omega^{12}) \end{aligned} \right\} \quad (22)$$

l'ensemble des conditions pour que la précision soit inchangée.

Mais les conditions supplémentaires suivantes :

$$\frac{\Omega^{12}}{r_1 r_2} = \frac{\Omega^{11} - r_1}{r_1^2} = \frac{\Omega^{22} - r_2}{r_2^2} \quad (23)$$

déterminent, avec (22), deux traitements confondus (ou à une distance nulle).

### 3.2. Un traitement est confondu avec un bloc

Soit  $i_0$  le traitement,  $j_0$  le bloc.

S'ils sont à une distance nulle :

$$0 = \sum_h (f_h(i_0) - f_h(j_0))^2 = \sum_h \left[ \lambda_h^{\frac{1}{2}} t_{hi_0} - \sum_i (a_{ij_0}/k_{j_0}) t_{hi} \right]^2 = 0.$$

Alors :

$$\lambda_h^{\frac{1}{2}} t_{hi_0} = \sum_i (a_{ij_0}/k_{j_0}) t_{hi}$$

et

$$[\lambda_h^{\frac{1}{2}} - (a_{i_0j_0}/k_{j_0})] t_{hi_0} = \sum_{i \neq i_0} (a_{ij_0}/k_{j_0}) t_{hi} \quad (24)$$

pour tout  $h$ .

D'autre part, l'équation (16) implique que :

$$\mathbf{TT}' = r^{-\delta} \quad (25)$$

alors :

$$\begin{aligned} \sum_h (\lambda_h^{\frac{1}{2}} - a_{i_0j_0}/k_{j_0}) t_{hi_0}^2 &= \sum_h \sum_{i \neq i_0} (a_{ij_0}/k_{j_0}) t_{hi_0} t_{hi} \\ \sum_h \lambda_h^{\frac{1}{2}} t_{hi_0}^2 &= a_{i_0j_0}/(r_{i_0} k_{j_0}) \end{aligned} \quad (26)$$

est la condition pour que la distance entre le bloc  $j_0$  et le traitement  $i_0$  soit nulle.

*Remarque.* — Dans le cas du plan J-équilibré,

$$\lambda_h = \lambda \quad \text{pour tout } h \leq t_0$$

et (26) est alors équivalente à

$$\lambda = a_{i_0j_0}/k_{j_0} \quad (27)$$

### 3.3. Exemple

Trouver les matrices  $\mathbf{\Omega}_{3 \times 3}$  telles que  $(x_1 + x_2 - 2x_3)$  et  $(3x_1 - 2x_2 - 2x_3)$  soient des  $r$ -contrastes, et que le plan soit J-équilibré.

$$\left. \begin{aligned} r_1 + r_2 - 2r_3 &= 0 \\ 3r_1 - 2r_2 - 2r_3 &= 0 \end{aligned} \right\} \Rightarrow r_1 = 6\alpha, \quad r_2 = 4\alpha, \quad r_3 = 5\alpha, \quad N = 15\alpha.$$

Prenons le premier  $r$ -contraste comme contraste de base, i. e. comme axe factoriel, et cherchons le second :

$$(1 \quad 1 \quad -2) \begin{pmatrix} 6 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} t_{21} \\ t_{22} \\ t_{23} \end{pmatrix} = 0 \Rightarrow \text{une solution}$$

$$t_{21} = 4, \quad t_{22} = -6, \quad t_{23} = 0$$

$$\mathbf{T} = (15\alpha)^{-\frac{1}{2}} (\mathbf{t}_1 \mathbf{t}_2 \mathbf{1})$$

$$\mathbf{T}'r^\delta\mathbf{T} = \mathbf{I} \Rightarrow \mathbf{t}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \quad \mathbf{t}_2 = \begin{pmatrix} 1 \\ -3/2 \\ 0 \end{pmatrix}$$

$$\mathbf{T} = \frac{1}{(15\alpha)^{\frac{1}{2}}} \begin{pmatrix} 1/\sqrt{2} & 1 & 1 \\ 1/\sqrt{2} & -3/2 & 1 \\ -2/\sqrt{2} & 0 & 1 \end{pmatrix}$$

$$\mathbf{T}(\mathbf{1} - \lambda)^{-\delta}\mathbf{T}' = \mathbf{\Omega} = [\alpha(\mathbf{1} - \lambda)]^{-1} \left[ \begin{pmatrix} 6 \\ 4 \\ 5 \end{pmatrix}^{-\delta} - (\lambda/15)\mathbf{11}' \right]$$

(Le plan est J-équilibré, si  $\lambda_1 = \lambda_2 \neq 0, \lambda_3 = 0$ ).

#### 4. APPLICATION AUX PLANS BINAIRES ÉQUILIBRÉS

Soit un BIE  $(t, b, r, k, \lambda)$  (\*) où l'on sait que  $\lambda = \frac{r(k-1)}{t-1}$  ( $\lambda$  est le nombre de fois où deux traitements se retrouvent dans un même bloc)

$$\begin{aligned} \mathbf{nn}' &= (r - \lambda)\mathbf{I} + \lambda\mathbf{11}' \\ \mathbf{\Omega}^{-1} &= r\mathbf{I} - k^{-1}\mathbf{nn}' + r^2\mathbf{N}^{-1}\mathbf{11}' \\ r(\mathbf{I} - (rk)^{-1}\mathbf{nn}' + r\mathbf{N}^{-1}\mathbf{11}') &= r(\mathbf{I} - \mathbf{H}) \\ \mathbf{H} &= (rk)^{-1}[(r - \lambda)\mathbf{I} + \lambda\mathbf{11}'] - r\mathbf{N}^{-1}\mathbf{11}' \\ \left. \begin{aligned} \mathbf{H}_{ii} &= k^{-1} - r\mathbf{N}^{-1} \\ \mathbf{H}_{ij} &= \lambda(rk)^{-1} - r\mathbf{N}^{-1} \quad i \neq j \end{aligned} \right\} i, j = 1, \dots, t. \end{aligned}$$

Cherchons les valeurs propres  $\lambda_h$  de  $\mathbf{H}$  :

$$\begin{aligned} \det |\mathbf{H} - x\mathbf{I}| &= [k^{-1} - r\mathbf{N}^{-1} - x + (t-1)(\lambda(rk)^{-1} - r\mathbf{N}^{-1})] \cdot [k^{-1} - x - \lambda(rk)^{-1}]^{t-1} \\ &= 0 \end{aligned}$$

a)  $x = 0$

b)  $x = k^{-1} - \lambda(rk)^{-1} = \boxed{\frac{r - \lambda}{rk} = \lambda_h} \quad h = 1, \dots, t - 1.$

---

(\*) La notation adoptée est celle que l'on peut trouver par exemple dans KENDALL et STUART (1967).

La somme des valeurs propres

$$\sum_h \lambda_h = (t-1) \frac{r-\lambda}{rk} = \frac{t-k}{k} \quad (28)$$

est une mesure du lien entre les blocs et les traitements.

Pour un BIE :  $t \geq k$  (l'égalité n'ayant lieu que si le plan est de plus orthogonal ;  $n = 11'$ ).

$$\text{D'autre part, } \mu_h = \frac{\lambda_h}{1-\lambda_h} = \frac{r-\lambda}{\lambda t}$$

et l'efficacité

$$\varepsilon = \frac{t}{t + \sum \mu_h} = \frac{\lambda t^2}{rkt - r + \lambda} = \frac{(k-1)t^2}{(k-1)t^2 + (t-1)(t-k)}$$

Un « bon » plan d'expérience est tel que  $\varepsilon$  soit proche de 1, ce qui ne peut être réalisé que si  $\sum \lambda_h$  est voisin de 0, c'est-à-dire si deux traitements quelconques sont « souvent » dans un même bloc ( $k$  voisin de  $t$ , taille des blocs voisine du nombre de traitements).

Cherchons  $\alpha = \alpha(k, t)$ , le coefficient de proportionnalité entre  $\varepsilon$  et

$$\sum_h \lambda_h :$$

$$\frac{(k-1)t^2}{(k-1)t^2 + (t-1)(t-k)} = \alpha \frac{t-k}{k}$$

$$\Rightarrow \alpha = \alpha(k, t) = \frac{k(k-1)t^2}{(t-k)(t^2(k-1) + (t-1)(t-k))}$$

Alors :

$$\text{si } k \leq \frac{t}{2} \quad \alpha < 1$$

$$\text{si } k \geq \frac{t+1}{2} \quad \alpha > 1$$

Il est souhaitable que  $\varepsilon > \sum_h \lambda_h$ ,

donc que  $\alpha > 1$ .

Un « bon » BIE doit donc vérifier

$$(t+1)/2 \leq k \leq t$$

(29)

### 5. COMPARAISON AVEC LA MÉTHODE DE KUIPER ET CORSTEN

5.1. Kuiper (1952) et Corsten (1958) ont proposé une méthode de calcul d'analyse de variance fondée sur des considérations essentiellement géométriques, ces auteurs considèrent deux applications, l'une de l'espace des traitements dans celui des blocs :

$$\begin{aligned}
 p : \mathbf{R}^t &\rightarrow \mathbf{R}^b \\
 \mathbf{x} &\mapsto \mathbf{k}^{-\delta} \mathbf{A}' \mathbf{x}
 \end{aligned}$$

et l'autre de l'espace des blocs dans celui des traitements :

$$\begin{aligned}
 q : \mathbf{R}^b &\rightarrow \mathbf{R}^t \\
 \mathbf{y} &\mapsto \mathbf{r}^{-\delta} \mathbf{A} \mathbf{y}.
 \end{aligned}$$

Les axes factoriels  $\mathbf{t}_h$  (respectivement  $\tilde{\mathbf{t}}_h$ ) sont des directions fixes pour l'application composée  $q \circ p$  (respectivement  $p \circ q$ ).

Alors, il suffit de remarquer que dans l'expression de  $\mathbf{\Omega} = \left( \mathbf{I} - \sum_{l=1}^{\infty} \mathbf{M}_0^l \right) \mathbf{r}^{-\delta}$ ,

l'application  $p \circ q$  est de matrice  $\mathbf{M}_0$  pour aboutir à une nouvelle expression

$$\boxed{\mathbf{\Omega} = \sum_{j=0}^{\infty} (q \circ p)^j(\mathbf{T}) \cdot \mathbf{T}'} = \mathbf{r}^{-\delta} + \sum_{j=1}^{\infty} (q \circ p)^j(\mathbf{T}) \cdot \mathbf{T}' \tag{30}$$

où les colonnes de  $\mathbf{T}$  sont les axes  $\mathbf{t}_h$  (multipliés par  $\mathbf{N}^{-\frac{1}{2}}$ ).

Cette égalité peut également être donnée sous une autre forme, plus intéressante d'un point de vue analytique :

$$\mathbf{\Omega}^{-1} = \mathbf{r}^{\delta} - \mathbf{A} \mathbf{k}^{-\delta} \mathbf{A}' = \mathbf{r}^{\delta} (\mathbf{I} - \mathbf{M}_0)$$

d'où

$$\boxed{\mathbf{\Omega} = \sum_{j=0}^{\infty} (q \circ p)^j(\mathbf{r}^{-\delta})} = \mathbf{r}^{-\delta} + \sum_{j=1}^{\infty} (q \circ p)^j(\mathbf{r}^{-\delta}) \tag{31}$$

expression identique à celle fournie par Calinski (1972).

La matrice de covariance  $\mathbf{\Omega}$  peut donc s'exprimer comme la somme de la matrice de covariance du plan orthogonal correspondant, et des images de cette dernière par les puissances de l'application composée  $q \circ p$ .



### 5.2. Estimateurs des paramètres traitement et bloc

Si dans le modèle linéaire général  $y = X\beta + e$  nous considérons une expérience en blocs avec un facteur à éliminer (l'hétérogénéité des parcelles) et un facteur d'intérêt (les traitements) nous nous intéressons généralement à l'estimation des composantes dans  $\beta$  des blocs, soit  $\beta_1$ , et des traitements, soit  $\beta_2$ . Si nous introduisons les quantités  $Q_2 = T_2 - nk^{-\delta}T_1$ ,  $Q_1 = T_1 - n'r^{-\delta}T_2$ , où  $T_1$  et  $T_2$  sont respectivement les vecteurs des totaux des observations des  $b$  blocs et des  $t$  traitements alors on sait que  $\hat{\beta}_2 = \Omega Q_2$ , d'où

$$\hat{\beta}_2 = \sum_{j=0}^{\infty} (q \circ p)^j (r^{-\delta} Q_2) \quad (32)$$

de même, en introduisant le total général  $G = 1'_t T_2 = 1'_b T_1$

$$\begin{aligned} \hat{\beta}_1 &= k^{-\delta} T_1 - k^{-\delta} n' \hat{\beta}_2 - (G/N) 1_b \\ &= k^{-\delta} T_1 - k^{-\delta} A' \hat{\beta}_2 - (G/N) 1_b = k^{-\delta} T_1 - p(\hat{\beta}_2) - (G/N) 1_b \end{aligned}$$

$$\hat{\beta}_1 = k^{-\delta} T_1 - (G/N) 1_b - \sum_{j=0}^{\infty} (p \circ q)^j \circ p (r^{-\delta} Q_2) \quad (33)$$

## 6. RELATIONS AVEC L'ANALYSE DES CORRÉLATIONS CANONIQUES

6.1. Soit, sous l'hypothèse du modèle linéaire additif,  $Q_2$  et  $Q_1$  les estimateurs (non ajustés) des traitements et des blocs, dont les expressions sont données ci-dessus.

Nous considérons maintenant les blocs et les traitements comme deux ensembles de variables entre lesquelles on cherche les corrélations canoniques, pour cela on sait que (Chakrabarti, 1962) :

$$\begin{aligned} \text{Var}(Q_2) &= \sigma^2 C_2 = \sigma^2 [r^\delta - nk^{-\delta} n'] \\ \text{Var}(Q_1) &= \sigma^2 C_1 = \sigma^2 [k^\delta - n'r^{-\delta} n] \\ \text{Cov}(Q_2, Q_1) &= -\sigma^2 [C_2 r^{-\delta} n] = -\sigma^2 [nk^{-\delta} C_1]. \end{aligned}$$

Les corrélations canoniques sont, par définition, les racines de l'équation :

$$\det(\Sigma(\rho)) = 0 \quad (34)$$

où

$$\Sigma(\rho) = \begin{pmatrix} \rho\mathbf{C}_2 & -\mathbf{C}_2\mathbf{r}^{-\delta}\mathbf{n} \\ -\mathbf{C}_1\mathbf{k}^{-\delta}\mathbf{n}' & \rho\mathbf{C}_1 \end{pmatrix}$$

$\Sigma(\rho)$  peut se factoriser de la manière suivante :

$$\Sigma(\rho) = \begin{pmatrix} \mathbf{C}_2 & 0 \\ 0 & \mathbf{C}_1 \end{pmatrix} \begin{pmatrix} \rho\mathbf{I} & -\mathbf{r}^{-\delta}\mathbf{n} \\ -\mathbf{k}^{-\delta}\mathbf{n}' & \rho\mathbf{I} \end{pmatrix}.$$

Les corrélations canoniques, solutions de (34), sont aussi solutions de

$$\det \begin{pmatrix} \rho\mathbf{I} & -\mathbf{r}^{-\delta}\mathbf{n} \\ -\mathbf{k}^{-\delta}\mathbf{n}' & \rho\mathbf{I} \end{pmatrix} = 0 \tag{35}$$

On peut encore montrer que les racines non nulles de (35) sont les mêmes que celles de

$$\det \begin{pmatrix} \rho\mathbf{I} & -\mathbf{r}^{-\delta}\mathbf{A} \\ -\mathbf{k}^{-\delta}\mathbf{A}' & \rho\mathbf{I} \end{pmatrix} = 0 \tag{36}$$

avec le même ordre de multiplicité (Kendall et Stuart, 1967).

De plus (36) est équivalent à

$$\begin{aligned} \det(\rho^2\mathbf{I} - \mathbf{M}_0) &= 0 \\ &= \det(\rho^2\mathbf{I} - \tilde{\mathbf{M}}_0) \end{aligned} \tag{37}$$

où  $\tilde{\mathbf{M}}_0 = \mathbf{k}^{-\delta}\mathbf{A}'\mathbf{r}^{-\delta}\mathbf{A}$

Les corrélations canoniques entre les estimateurs des blocs et des traitements sont donc les racines carrées des valeurs propres  $\lambda_h$  de  $\mathbf{M}_0$  ou  $\tilde{\mathbf{M}}_0$ . Les combinaisons linéaires associées sont par conséquent

$$t'_h\mathbf{Q}_2 \quad \text{et} \quad \tilde{t}'_h\mathbf{Q}_1 \quad (h = 1, \dots, t_0 = rg\mathbf{M}_0).$$

De plus, le nombre E, intervenant dans le calcul de l'efficacité, vaut

$$E = \sum_{h=1}^{t_0} \frac{\rho_h^2}{1 - \rho_h^2}.$$

Il peut s'interpréter comme une mesure de la part de variation des traitements « expliquée » par les blocs, relativement à la part de variation indépendante. Le plan d'expérience est d'autant plus efficace que E est petit.

## 6.2. Pour une présentation plus géométrique

Avant de présenter deux exemples numériques, nous tenons à justifier la formulation matricielle que nous avons adoptée et qui se révèle assez lourde à l'usage.; nous l'avons tout de même suivie pour deux raisons : la première est que notre présentation est surtout destinée aux spécialistes de l'analyse de la variance habitués par les ouvrages de langue anglaise à cette notation. La seconde est qu'elle se prête bien au traitement numérique ; elle est très voisine des procédures algorithmiques utilisables sur ordinateur qu'elles soient itératives et rapidement convergentes (Verdooren, 1969), ou qu'elles fassent appel à des séquences finies d'opérations de « balayage » sur le vecteur observation défini au paragraphe 5.2 (Wilkinson, 1970). Enfin, il est facile dans cette voie d'étendre les résultats au cas de dispositifs à plus de deux facteurs en introduisant la notion de correspondances multiples (Kettenring, 1971 ; Lebart, 1975).

Néanmoins, nous reconnaissons bien volontiers qu'une présentation géométrique telle qu'elle est faite par Caillez et Pagès (1976) aurait allégé notre exposé ; le rôle des multiplications par  $r^\delta$  et  $k^\delta$  correspondrait alors à un simple choix de métrique. Cette approche ne semble toutefois pas encore entrée dans les habitudes des statisticiens qui n'y font référence qu'à mots couverts (Herr, 1976).

## 7. EXEMPLES

Les deux exemples proposés sont extraits de Pearce (1963), ils ont été traités sur un programme classique d'analyse factorielle (Roux, 1975).

### 7.1. Plan de type S ( $t = 4$ , $b = 6$ )

Un plan de type S est formé d'un sous-plan totalement équilibré avec un traitement supplémentaire. Ici, le traitement supplémentaire, noté 0,

#### Légende du tableau I

Résultats numériques de l'analyse factorielle

a) Matrice d'incidence au plan

b) Valeurs propres de la matrice  $V$

c) Résultats numériques liés aux vecteurs propres  $c_1$  pour les traitements,  $c_2$  pour les blocs ;  $1 \neq F$ ,  $2 \neq F$  projection sur le premier facteur, sur le deuxième...

$CTR_i$  : contribution relative à l'axe factoriel  $i$

$CTA_i$  : contribution absolue à l'axe factoriel  $i$

$QLT = \sum_i CTR_i$ , qualité de l'ajustement au modèle factoriel

$POID = r_i/N$ .

TABLEAU 1. — *Plan de type S* (S. C. PEARCE, 1963)

a)

Nom J(J)	B1	B2	B3	B4	B5	B6		
$k \rightarrow$	4	4	4	4	4	4	24	= N
O . . . . .	2	2	2	2	2	2	12	
A . . . . .	1	0	0	1	1	1	4	
B . . . . .	0	1	1	1	0	1	4	
C . . . . .	1	1	1	0	1	0	4	

↑  
r

b)

Num	Iter	Val propre	Pourcent	Cumul
2	0	0.1249998	50.000	50.000
3	3	0.1249992	50.000	100.000
4	1	0.0000006	0.000	100.000
5	4	0.0000000	0.000	100.000
6	1	- .0000003	- 0.000	100.000

c<sub>1</sub>)

II	QLT	POID	INR	1 # F	CTR	CTA	2 # F	CTR	CTA
O . . . . .	0	5 000	0	0	0	0	0	0	0
A . . . . .	1 000	1 667	333	67	9	6	- 703	991	661
B . . . . .	1 000	1 667	333	- 642	827	551	294	173	115
C . . . . .	1 000	1 667	333	576	664	443	410	336	224
		24.0	1 000			1 000			1 000

c<sub>2</sub>)

J1	QLT	POID	INR	1 # F	CTR	CTA	2 # F	CTR	CTA
B1 . . . . .	1 000	1 667	167	455	827	276	- 207	173	58
B2 . . . . .	1 000	1 667	167	- 46	9	3	498	991	330
B3 . . . . .	1 000	1 667	167	- 46	9	3	498	991	330
B4 . . . . .	1 000	1 667	167	- 406	664	221	- 289	336	112
B5 . . . . .	1 000	1 667	167	455	827	276	- 207	173	58
B6 . . . . .	1 000	1 667	167	- 406	664	221	- 289	336	112
		24.0	1 000			1 000			1 000

est orthogonal aux blocs (ces plans ont été traités de façon générale par Calinski et Ceranka (1974)). La matrice  $M_0$  n'est donc pas de rang maximum, soit  $t - 1 = 3$ , mais de rang 2. Les deux valeurs propres non nulles identiques correspondent au sous-plan formé des traitements A, B et C. Cette valeur propre vaut  $1/8$ .

$$E = \sum_{h=1}^{rgM_0} \lambda_h / (1 - \lambda_h) = 2 \cdot \frac{1/8}{7/8} = 2/7$$

et l'efficacité  $\varepsilon = \frac{t}{t + E} = 14/15$ .

Les projections des traitements sur le plan des deux premiers axes factoriels mettent en évidence des contrastes entre A, B et C. Le traitement O est projeté à l'origine à cause de l'orthogonalité. Les propriétés d'équilibre du plan se traduisent sur la figure par des symétries. Les blocs deux à deux confondus forment un triangle équilatéral tel que la distance au traitement non représenté dans les deux blocs considérés soit maximum. Tous ces résultats sont clairement représentés sur la figure 1.

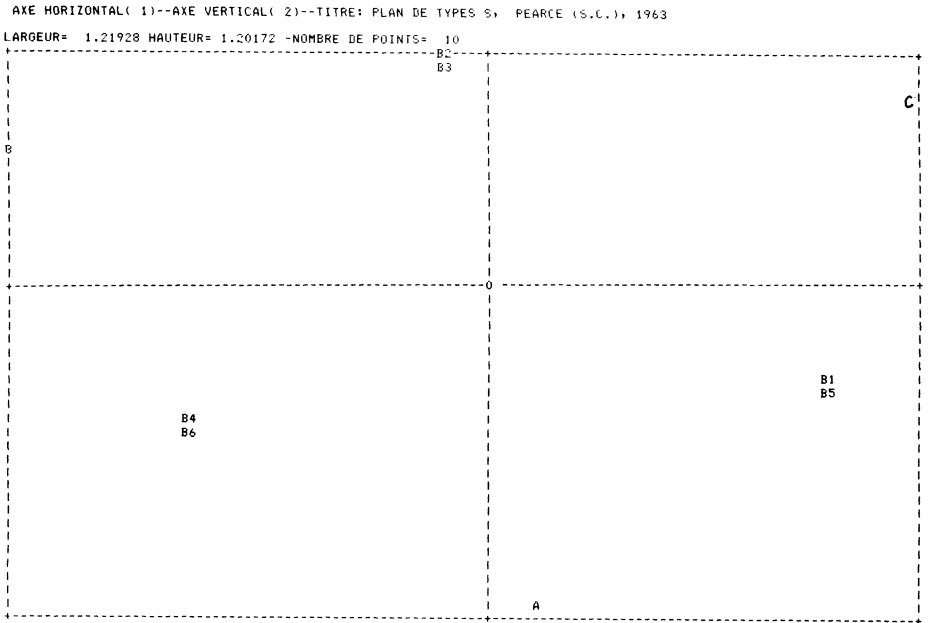


FIG. 1. — Plan de type S.  
 Projection des blocs et des traitements sur le plan des facteurs un et deux.

7.2. Plan de type F ( $t = b = 6$ )

Un plan de type F est factorisable en sous-plans totalement équilibrés. Comme les traitements sont équirépétés, les  $r$ -contrastes sont des contrastes simples et l'examen des projections des traitements sur les facteurs montrent deux groupes (A, B, C) et (D, E, F) formant des contrastes indépendamment l'un de l'autre. Remarquons que les blocs ont les mêmes propriétés (fig. 2 a et 2 b).

Les valeurs propres, d'ordre 2 chacune, valent  $3/64$  et  $1/64$

$$\varepsilon = 61/62 .$$

Ce plan a donc une très bonne efficacité.

Les corrélations canoniques

$$0,22 ; 0,22 ; 0,125 \text{ et } 0,125$$

entre estimateurs des blocs et des traitements sont donc relativement faibles.

TABLEAU 2. — Plan de type F (S. C. PEARCE, 1963)

a)

Nom J(J)	B1	B2	B3	B4	B5	B6		
$k \rightarrow$	8	8	8	8	8	8	48	= N
A . . . . .	2	1	1	1	1	2	8	
B . . . . .	1	2	1	2	1	1	8	
C . . . . .	1	1	2	1	2	1	8	
D . . . . .	1	1	2	2	1	1	8	
E . . . . .	2	1	1	1	2	1	8	
F . . . . .	1	2	1	1	1	2	8	

↑  
 $r$

b)

Num	Iter	Val propre	Pourcent	Cumul
2	1	0.0468751	37.500	37.500
3	4	0.0468748	37.500	75.000
4	2	0.0156250	12.500	87.500
5	1	0.0156246	12.500	100.000
6	1	0.0000001	0.000	100.000

(Même légende qu'au tableau I)

TABLEAU 2 (suite).

c<sub>1</sub>)

II	QLT	POID	INR	1 # F	CTR	CTA	2 # F	CTR	CTA	3 # F	CTR	CTA	4 # F	CTR	CTA	
A	1 000	1 667	167	287	658	292	107	92	41	-66	36	48	-163	214	285	
B	1 000	1 667	167	-235	446	198	195	304	135	175	246	327	24	4	6	
C	1 000	1 667	167	-50	20	9	-301	730	324	-107	93	125	140	157	209	
D	1 000	1 667	167	-286	658	292	-106	92	41	-66	36	48	-163	214	285	
E	1 000	1 667	167	236	446	198	-194	304	135	175	246	327	24	4	6	
F	1 000	1 667	167	51	20	9	302	730	324	-107	93	125	140	157	209	
		48.0	1 000			1 000			1 000			1 000				1 000

c<sub>2</sub>)

J1	QLT	POID	INR	1 # F	CTR	CTA	2 # F	CTR	CTA	3 # F	CTR	CTA	4 # F	CTR	CTA	
B1	1 000	1 667	167	302	730	324	-50	20	9	108	93	125	-139	157	209	
B2	1 000	1 667	167	-106	92	41	287	658	292	67	36	48	164	214	285	
B3	1 000	1 667	167	-194	304	135	-235	446	198	-174	246	327	-23	4	6	
B4	1 000	1 667	167	-301	730	324	51	20	9	108	93	125	-139	157	209	
B5	1 000	1 667	167	107	92	41	-286	658	292	67	36	48	164	214	285	
B6	1 000	1 667	167	195	304	135	236	446	198	-174	246	327	-23	4	6	
		48.0	1 000			1 000			1 000			1 000				1 000

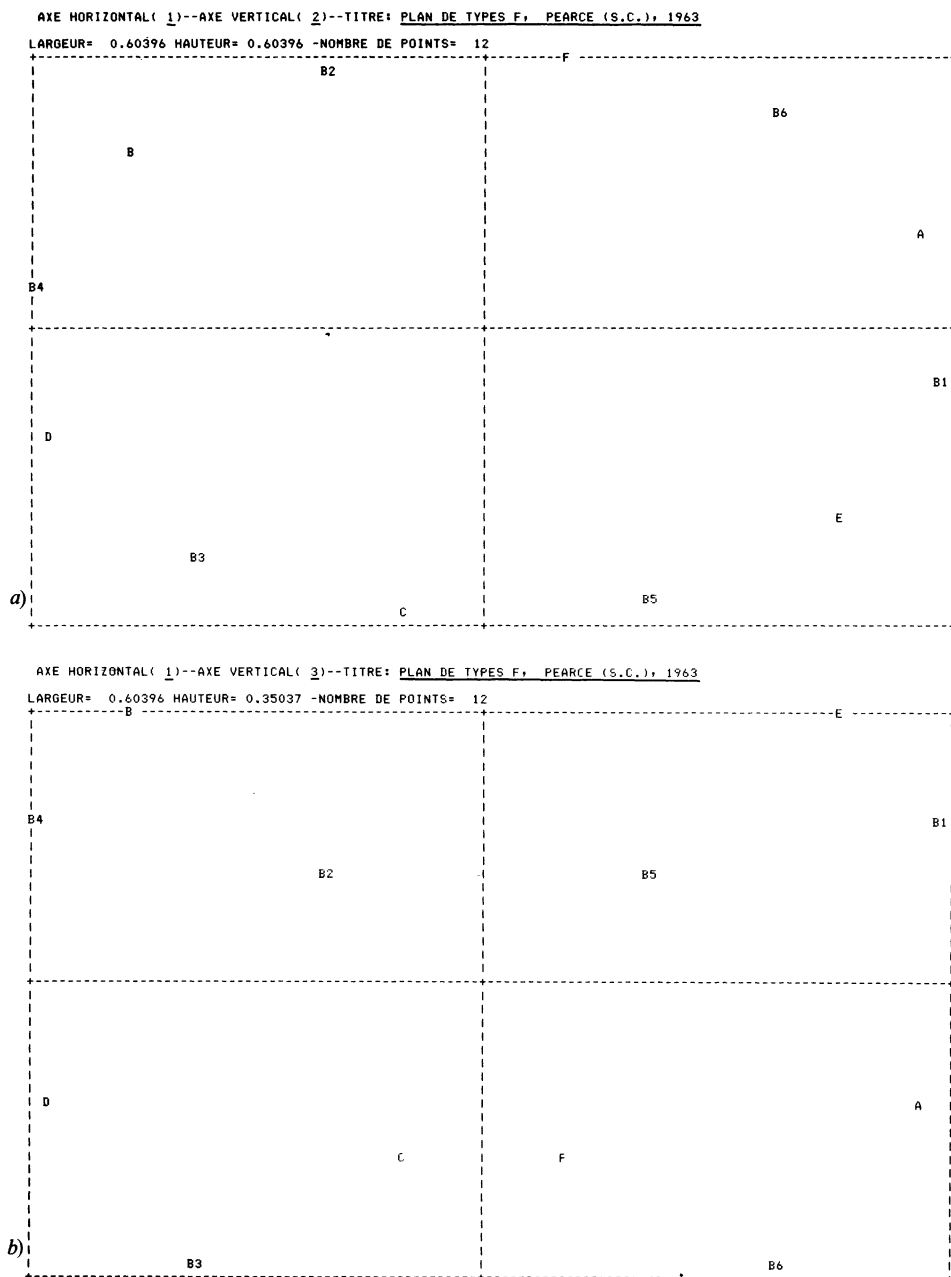


FIG. 2. — Plan de type F.

- a) Projection des blocs et des traitements sur le plan des facteurs un et deux.
- b) Projection des blocs et des traitements sur le plan des facteurs un et trois.



On voit donc sur ces deux exemples comment une analyse factorielle permet d'analyser très rapidement la structure d'un plan d'expérience.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] J. P. BENZÉCRI, *L'analyse des correspondances*. Dunod, Paris, 1973, p. 628.
- [2] F. CAILLEZ, J. P. PAGÈS, Introduction à l'analyse des données *S. M. A. S. H.*, Paris, 1976, p. 640.
- [3] T. CALINSKI, On some desirable patterns in block designs. *Biometrics*, t. 27, 1971, p. 275-292.
- [4] T. CALINSKI, On non orthogonal experiments, *Biom. Z.*, t. 14, 2, 1972, p. 73-84.
- [5] T. CALINSKI, B. CERANKA, Supplemented block designs. *Biom. Z.*, t. 16, 5, 1974, p. 299-305.
- [6] M. C. CHAKRABARTI, *Mathematics of design and analysis of experiments*. Asia Pub. House, Bombay, 1962, p. 126.
- [7] W. G. COCHRAN, G. M. COX, *Experimental designs*. Wiley, New York, 1968, p. 618.
- [8] L. A. CORSTEN, Vectors, a tool in statistical regression theory. *Medelingen van de Landbouwhogeschool te Wageningen*, t. 58, 1958, p. 1-92.
- [9] D. G. HERR, A geometric characterisation of connectedness in a two way design. *Biometrika*, t. 63, 1, 1976, p. 93-100.
- [10] M. O. HILL, Correspondence analysis: a neglected multivariate method. *Appl. Stat.*, t. 23, 3, 1974, p. 340-354.
- [11] R. M. JONES, On a property of incomplete blocks. *J. R. Statist. Soc. B*, t. 21, 1959, p. 172-179.
- [12] O. KEMPTHORNE, *The design and analysis of experiments*. Wiley, New York, 1952, p. 632.
- [13] M. G. KENDALL, A. STUART, *The advanced theory of statistics*, Volume 2, Griffin, Londres, 1967, p. 690.
- [14] J. R. KETTENRING, Canonical analysis of several sets of variables. *Biometrika*, t. 58, 3, 1971, p. 433-450.
- [15] N. H. KUIPER, Variantie-analyse. *Statistica*, t. 6, 1952, p. 149-194.
- [16] L. LEBART, L'orientation du dépouillement de certaines enquêtes par l'analyse des correspondances multiples. *Consommation*, t. 2, 1975, p. 73-96.
- [17] L. LEBART, J. P. FÉNELON, *Informatique et Statistique Appliquée*. Dunod, Paris, 1975, p. 440.
- [18] S. C. PEARCE, The use and classification of non orthogonal designs. *J. R. Statist. Soc.*, t. 126, A, 1963, p. 353-377.
- [19] S. C. PEARCE, The efficiency of block designs in general. *Biometrika*, t. 57, 2, 1970, p. 339-346.
- [20] S. C. PEARCE, T. CALINSKI, T. F. MARSHALL, The basic contrasts of an experimental design with special reference to the analysis of data. *Biometrika*, t. 61, 3, 1974, p. 449-460.
- [21] C. ROUX, *Analyse factorielle des correspondances*. C. N. R. Z.-Biom., Jouy-en-Josas, 75/05, 1975, p. 17.
- [22] L. R. VERDOOREN, Representation of ANOVA models with vectors and vector space. *Statistica Neerlandica*, t. 23, 1969, p. 53-65.
- [23] G. N. WILKINSON, A general recursive procedure for analysis of variance. *Biometrika*, t. 57, 1970, p. 19-46.

(Manuscrit reçu le 3 juin 1976)